



## **The Hmong Diaspora: Preserved South-East Asian genetic ancestry in French Guianese Asians.**

Nicolas Brucato, Stéphane Mazières, Evelyne Guitard, Pierre-Henri Giscard, Etienne Bois, Georges Larrouy, Jean-Michel Dugoujon

### **► To cite this version:**

Nicolas Brucato, Stéphane Mazières, Evelyne Guitard, Pierre-Henri Giscard, Etienne Bois, et al.. The Hmong Diaspora: Preserved South-East Asian genetic ancestry in French Guianese Asians.. *Comptes Rendus Biologies*, 2012, 335 (10-11), pp.698-707. 10.1016/j.crv.2012.10.003 . hal-00762359

**HAL Id: hal-00762359**

**<https://hal.science/hal-00762359>**

Submitted on 7 Dec 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# **The Hmong Diaspora: Preserved South-East Asian genetic ancestry in French Guianese Asians**

## **La Diaspora Hmong: un patrimoine génétique conservé chez la population asiatique de Guyane française**

Nicolas Brucato,<sup>1,2</sup> Stéphane Mazières,<sup>3§</sup> Evelyne Guitard,<sup>1</sup> Pierre-Henri Giscard,<sup>4</sup> Etienne Bois,<sup>†</sup> Georges Larrouy,<sup>1</sup> Jean-Michel Dugoujon<sup>1</sup>

<sup>1</sup> UMR5288 CNRS, Laboratoire d'Anthropologie Moléculaire et Imagerie de Synthèse (AMIS), Université Paul Sabatier Toulouse III, Toulouse, France.

<sup>2</sup> Language and Genetics Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.

<sup>3</sup> Aix Marseille Université, CNRS, EFS-AM, ADES UMR 7268, 13344, Marseille, France

<sup>4</sup> Institut des Déserts et des Steppes – Muséum, Paris, France.

§ Corresponding author:

UMR 7268 ADES Anthropologie, Droit, Ethique et Santé

Aix-Marseille Université, CNRS, EFS-AM

Faculté de Médecine - Secteur Nord- Bâtiment A - CS80011

51, Bd Pierre Dramard

13344 MARSEILLE Cedex 15 - FRANCE

[stephane.mazieres@univ-amu.fr](mailto:stephane.mazieres@univ-amu.fr)

## **ABSTRACT**

The Hmong Diaspora is one of the widest modern human migrations. Mainly localised in South-East Asia, the United States of America, and metropolitan France, a small community has also settled the Amazonian forest of French Guiana. We have biologically analysed 62 individuals of this unique Guianese population through three complementary genetic markers: mitochondrial DNA (HVS-I/II and coding region SNPs), Y-chromosome (SNPs and STRs), and the Gm allotypic system. All genetic systems showed a high conservation of the Asian gene pool (Asian ancestry: mtDNA=100.0%; NRY=99.1%; Gm=96.6%), without a trace of founder effect. When compared across various Asian populations, the highest correlations were observed with Hmong-Mien groups still living in South-East Asia ( $F_{st} < 0.05$ ;  $p$ -value  $< 0.05$ ). Despite a long history punctuated by exodus, the French Guianese Hmong have maintained their original genetic diversity.

Keywords: Anthropology, genetic markers, Asian populations, gene diversity, Diaspora

Mots-clés: Anthropologie, marqueurs génétiques, populations asiatiques, diversité génétique, Diaspora

## INTRODUCTION

The Hmong Diaspora is widely scattered in Asia, but also in America, Europe and Oceania. This human migration would have originated from the South Chinese province of Kweichow [1]. Associated with Neolithic cultures of Daxi (5300-6400 BP) and Qujialing (4600-5000 BP), the Hmong is an ancient and an important component of the human diversity of South East Asia [2]. The Han expansion, during the 16<sup>th</sup> century, forced the Hmong to leave their heartland and to settle the Yunnan highlands [2]. There they developed an economy based on a shifting agriculture of crops like maize and millet, together with husbandry and hunting. These strategies permitted the Hmong to prosper among the other ethnic minorities of China, where they are also known as Miao. The Hmong share common cultural traits particularly with the Mien, including language, and genetic characteristics that cluster the Hmong and Mien together in the so-called Hmong-Mien group [3, 4].

Following a period of calm, since the 18<sup>th</sup> century the Hmong have conflicted at various times with the Chinese government. They took refuge in northern Indochina where their skills in shifted cultivation rapidly integrated them into the regional economic network, and progressively gave the Hmong a relative institutional power, under the French colonial influence. However, the downside of the strategic role of the Hmong was the direct implication in the First Indochina War (1946-1954), the Vietnam War (1959-1975) and the Secret War in Laos (1964-1975). The latter triggered a massive exodus of Hmong from Laos toward northern Thailand, but also toward other continents. Almost 200,000 individuals went into exile to the United States and 15,000 to metropolitan France. Most of these migrants preferred to live in occidental cities but others were steered toward an unusual destination: French Guiana[5] [5].

This South American territory, localised between Surinam and Brazil, has been a French colony since the 17<sup>th</sup> century. Today it only has 3 individuals per square kilometre.

Welcoming the Hmong onto its soil represented a double advantage for the French government: to shelter former cooperating people along with populating areas of low population density where conditions of life are similar to those found in Laos [5].

Approximately 2100 Hmong are living in French Guiana. They reinstated their former rural farming communities with presenting all the social characteristics of South East Asia, and preserved their original Hmong cultural identity. But more than an ethnic isolate, they are also completely integrated in the French Guianese social network by significantly increasing the local agricultural production. Today, they constitute an important component of the ethnic diversity of French Guiana that is primarily composed by Amerindians, Europeans, Creoles and Noir Marron populations [6-10].

The French Guianese example of the Hmong Diaspora represents a unique human migration that has never been studied genetically. To this aim, three different and complementary genetic systems were analysed in this study. Uniparental lineages were determined through the analysis of mitochondrial DNA (mtDNA) and non-recombinant Y chromosome (NRY) haplotypes. Through the large number of populations that have previously been genotyped for these two systems, informative data are available to identify the ancestry of each haplotype observed in this study [11, 12]. The third system was the biparental Gm allotypic system. This system comprises polymorphic antigenic markers located on the constant regions of the heavy chains of three subclasses of immunoglobulin G (IgG1, IgG2, and IgG3) [13, 14]. The Gm system has been frequently studied in human populations and has shown heterogeneous haplotype frequencies among populations [15, 16]. Our study proposes the first genetic insight into the Hmong Diaspora and its particular example in French Guiana. The present study will aim to answer two questions in particular: (i) Has the original Hmong genetic pattern been shaped by founder effect and admixture during their long exodus? (ii) Is the French Guianese Hmong (FGH) gene pool still close to the Hmong pool seen in South East Asia today?

## **MATERIALS AND METHODS**

### **Population Sampled**

A total of 141 individuals were sampled in the Cacao and Javouhey communities (Figure 1) in 1980 under the auspices of the Institut National de la Santé et de la Recherche Médicale (Paris, France). All samples were obtained with the informed consent of the participants. DNA was extracted recently by a phenol-chloroform protocol from sera, and stored at  $-20^{\circ}\text{C}$ . During the Cacao and Javouhey sample collection, pedigrees were recorded and used to selected 62 maternal and 33 paternal unrelated lineages for the present study.

### **Laboratory methods**

#### *mtDNA*

Maternal lineages were characterised by the sequencing of a large part of the D-loop region (16012-263) following previously described protocol [17]. All data were obtained on an ABI PRISM 3730 sequencer (PE, Applied Biosystems) and analysed with Sequence Scanner v.1.0 (PE, Applied Biosystems). For preliminary haplogroup assignment, each sequence was firstly aligned with rCRS [18] using BioEdit v.7.0.9.0, then tested for relevant single nucleotide polymorphisms (SNPs) of the coding region (positions 10400, 10398, 7028, 5178, 13263, 6392) typed by Restriction Fragment Length Polymorphisms with the following enzymes: *AluI*, *HpaI*, and *DdeI*. The final haplogroup assignment was obtained from the differences with the rCRS by the most recent mtDNA phylogeny [12, 19], then the haplotypes presented in relation to both the rCRS and RSRS references sequences [20].

#### *NR1*

Paternal lineages were determined by two types of markers. NR1 haplotypes are a combination of seventeen Short Tandem Repeats (STRs), typed using the AmpFLSTR

Yfiler® kit (PE, Applied Biosystems), and informative SNPs (UEPs: SRY 10831, M213, M9, M70, M22, Tat, 92R7, M173, P25, M122, M134, M175, P31, M50, M101, M119, SRY465, 47z, M88, M95, M216, M174, P197), typed by SNaPshot® minisequencing (PE, Applied Biosystems) using already published primers [11]. All data were read on an ABI PRISM 3730 sequencer and analysed with Genemapper v.4.0 (PE, Applied Biosystems). The YAP analysis has been obtained following Hammer and Horai [21]. The haplogroup assignment follows the recently updated NRY phylogeny [11] [22].

### *Gm*

The plasma samples were tested for G1m (1, 2, 3, 17), G2m (23), G3m (5, 6, 10, 11, 13, 14, 15, 16, 21, 24) immunoglobulin allotypic markers by using a classical haemagglutination inhibition method [23]. The reagents used are described in [24] and [25]. The assignment of Gm haplotypes follows the International System for Human Gene Nomenclature (ISGN). They are indicated in the following order in the text: G3m, G1m, and G2m. They are set in simplified forms [13]. Allotypes are separated by commas, and the subclasses are separated by semicolons. The notation “..” indicates that the sample was tested for G2m (23) but found to be negative, and “\*” stands for 10,11,13,14.

### **Data analyses**

Rates of genetic diversity based on mtDNA, NRY and Gm data were calculated using the ARLEQUIN 3.11 software package [26]. For each of these systems, a database was compiled from published studies and divided in eight groups: French Guiana, East Asia (China, Mongolia, Taiwan, Nepal, South Korea and Japan), mainland South East Asia (Thailand, Vietnam and Malaysia), South Asia (India, Sri Lanka, Bangladesh), Insular South East Asia (Indonesia, Philippines), West Europe (France, Italy, United Kingdom, Germany and Spain) (Additional file 1). The mtDNA database is composed by 139 populations representing 11 944

HVS-I associated with their corresponding haplogroup assignment. The NRY database was composed of 50 populations representing 11 094 individuals typed for Y-STR profiles. The Gm database contains information on 21 798 individuals from 127 populations (Additional file 1).

Haplotype networks were generated for two pan-Asiatic lineages, the mtDNA haplogroup M\* and the Y-chromosome haplogroup O3\*, using respectively the HVS-I data and the Y-STR core haplotype (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393) from the FGH data and all Asian comparable data, via the median-joining algorithm of Network v.4.5.1.6 ([www.fluxus-engineering.com](http://www.fluxus-engineering.com)). To obtain the most parsimonious networks the reticulation permissivity was set to zero. Data were pre-processed using the star contraction option in Network v.4.5.1.6 [27]. For the mtDNA data, hypermutable sites were identified by post-processing using the Steiner (MP) algorithm and removed from the analysis [28]. Weight of each Y-STR loci were characterised according to its variance in the O3a\* sample, as previously described [29]. Tajima's D and Fu's Fs tests were calculated with mtDNA data using the ARLEQUIN 3.11 software package [26].

Cross-population comparisons based on HVS-I mtDNAs, Y-STR haplotypes and Gm lineages were performed using ARLEQUIN 3.11 [26] and depicted through the Fst parameter. Significance is given for p-values under 0.05. Matrices of the Fst genetic distances and p-values are available upon request. All results obtained for the comparison between the FGH and each population of the database were graphically plotted on a map using Surfer v.8.0, using the location of each population given in the corresponding study. Percentage of homology (perfect match) between FGH data and the databases has been performed from HVS-I mtDNA and Y-STR core haplotype. Admixture estimates were calculated by two different methods: *mY* calculated by Admix v.2.0 from haplogroup assignments [30] and the Lineages Sharing rate (LS) from haplotypic data [31]. All estimates were computed from



10,000 bootstraps replications without taking into account the molecular distances. Five hypothetical parental populations were tested: Mainland South East Asia, Insular South East Asia, East Asia, South Asia and Europe. Each of these groups is composed by all corresponding data compiled in the database. A haplotype matching probability has been calculated for each FGH haplotypes in each of these five hypothetical parental populations.

## RESULTS

### mtDNA

#### *Gene Diversity*

A total of 33 different mtDNA haplotypes were characterised among the 62 Hmong individuals sampled in French Guiana (Additional file 2). The mtDNA sequence diversity computed ( $H_{\text{mtDNA}}=0.910 \pm 0.005$ ;  $MPD_{\text{mtDNA}}= 9.10 \pm 4.29$ ) was within the range generally observed in South East Asian populations [32, 33]. It is close to the ones present in Hmong-Mien populations living between the South of China and the North of Thailand [4, 34]. Although the Tajima's test is not significant ( $D = -1.298 \pm 0.075$ ), this diversity is associated with a significant value for Fu's test, revealing a population expansion ( $F_s = -16.276 \pm 0.000$ ). All mtDNA haplotypes were phylogenetically identified following the latest classification scheme [12]. The majority were found to belong to haplogroups frequent in Asia [35]. Among them, most were clustered in the M\* haplogroup (26.9%) revealing the importance of the Eurasian founder node in the complexity of the genetic structure in East Asia, particularly in the region interfacing India and China [36, 37]. The phylogenetic tree for this haplogroup showed no single FGH founder haplotype, but there is a presence of many haplotypes from within the overall Asian diversity (Additional File 3). The percentage of homology show that FGH M\* HVS-I haplotypes are shared at 70% by South Asian individuals (Additional File 6). Other haplogroups frequent in Asia were also found to be present, such as the B subclades

B5/B5a (22.5%), and B4 (3.2%), F1a (11.2%), the M subclades M5 (7.9%), M74a (3.2%), M8/C\* (3.2%), M25 and M7b (1.6% each), C5 (4.8%), N9a (4.8%), D\* (3.2%), and F3a1 (3.2%). Two individuals were classified within the H1C4b haplogroup frequent in Europe, but their HVS-I profile was homologous to those found in South Asia (Additional File 6). No mitochondrial haplogroup frequent in Amerindian or African groups was detected.

#### *Population cross-comparisons*

From these results, population pairwise  $F_{st}$  comparisons were computed from haplotypic data to evaluate the genetic distances between the FGH and each population present in the Asian database (Figure 2). Relatively low genetic divergences were observed ( $0 < F_{st} < 0.25$ ) with all the populations of the database, revealing the complexity of the Asian maternal diversity [37]. Twenty-four Asian populations presented significantly low genetic divergences ( $F_{st} < 0.05$ ;  $p$ -value  $< 0.05$ ) which identified the clearest links with the FGH: 13 from India, 6 from the South of China, 1 from Thailand, 1 from Vietnam, 1 from Sri Lanka, 1 from Mongolia and 1 from Bangladesh.

#### *Admixture analysis*

To estimate the potential contribution of geographical dispersal to the mitochondrial inheritance, admixture ratios were calculated according to different models from five hypothetical parental populations (Table 1). All estimators pinpoint South Asian ( $LS=0.453$ ;  $mY=0.711$ ; Table 1) and Mainland South East Asian ( $LS=0.201$ ;  $mY=0.289$ ; Table 1) populations as the major sources of the FGH maternal inheritance.

### **NR1**

#### *Gene Diversity*

Twenty four different NR1 haplotypes were detected among the 33 sampled FGH men (YHRD accession Number YA003693; Additional file 4). The genetic diversity revealed by

the Y-STR haplotypes is as high as the diversity found in East Asian populations ( $H_{NRY}=0.9773 \pm 0.0132$ ) [34]. These Y-STR haplotypes were phylogenetically identified by Y-SNP following the latest classification scheme [11]. They all belong to Asian haplogroups [38]: O3a (48.5%), D (24.2%), C (15.1%), O2a1 (6.1%) and O1a (6.1%). No FGH founder haplotype was detected within the phylogenetic tree of O3a\*. Instead, the FGH profiles were found to be broadly scattered within the East Asian diversity (Additional File 6). The percentage of homology rose to 78% with O3\* Y-STR core haplotypes found in East Asia (Additional File 6). No European, Amerindian or African NRY haplogroup was identified.

#### *Population cross-comparisons*

The population pairwise  $F_{st}$  comparisons performed for the core Y-STR haplotypes of the FGH with the Asian populations of the database (Figure 3) showed a higher geographical structuring than the maternal data. Although relatively low genetic divergences are observed ( $0.05 < F_{st} < 0.25$ ) with all the populations of the database, only three of them present significantly low genetic divergence ( $F_{st} < 0.05$ ). Furthermore, they are all located in a region covering the South of China and the North of Thailand.

#### *Admixture analysis*

Although the paternal genetic relationship appears to be localised, admixture tests were performed to estimate the ratio of the most probable parental populations (Table 2). All estimators localize major ancestries in East Asia ( $LS=0.459$ ;  $mY=1.000$ ; Table 2) and in Mainland South East Asia ( $LS=0.290$ ;  $mY=0.000$ ; Table 2).

### **Gm**

#### *Gene Diversity*

Among the 62 FGH typed, the observed phenotypes can be explained by four different Gm haplotypes. The inherent genetic diversity appears to be low but still within the range

observed in South East Asian populations ( $0.4424 \pm 0.0152$ ) [16]. Although Gm 21\*;1,17;.. (22.1%) and Gm 21\*;1,2,17;.. (3.4%) are as common in Asia as in Europe, haplotypes Gm 5\*;1,3;.. (71.6%) and Gm 15,16\*;1,17;.. (2.9%) are particularly present in Asian groups.

#### *Population cross-comparisons*

As expected for a biparental marker, the results obtained for the population pairwise  $F_{st}$  comparison shows an intermediate pattern to the uniparental markers (Figure 4). Although high genetic divergence was observed with the Tibetan groups ( $F_{st} > 0.25$ ), 20 of the 39 populations of the Asian database present relatively low  $F_{st}$  levels ( $0 < F_{st} < 0.25$ ), identifying the relative homogeneity of this system in East Asia [16, 39]. Nine populations have significantly low genetic divergence from the FGH ( $F_{st} < 0.05$ ), all located in the South of China or in the North of Thailand. These results converge with both the mtDNA and NRY analyses.

#### *Admixture analysis*

As for the uniparental markers, the contribution of hypothetical parental populations was estimated. Mainland South East Asia ( $LS=0.324$ ;  $mY=0.594$ ; Table 3) and Insular South East Asia ( $LS=0.288$ ;  $mY=0.363$ ; Table 3) are indicated as the major source of the FGH Gm inheritance.

## **DISCUSSION**

### **An Asian genetic diversity in French Guiana**

Herein we provided immunoglobulin, mtDNA and Y-chromosome data in two Hmong populations from French Guiana to define the dynamics of the Hmong settlement in the area. The combined use of uni- and biparental genetic markers is a crucial strategy to highlight the demographical patterns of modern human expansion, migration rates, co-evolution of genes

with cultural traits, and sex-biased admixture [40-45]. Previously, such procedure has furnished interesting insights about the genetic history of the understudied area [8, 10]. Although the Hmong arrived in French Guiana as recently as 1975, the Hmong Diaspora started centuries ago. During this complex history, the Hmong population rapidly adapted itself to each new environment, including in inhospitable areas. Two traits are particularly noticeable in aspects of the FGH gene pool assessed in French Guiana. Firstly, for many communities of Diasporas, the original genetic diversity is often shaped by founding effects and/or genetic drift, particularly when settling isolated locations [46]. This pattern could have been detected in the FGH, especially considering the small number of initial immigrants (around 2,000 individuals). However, on the contrary, the analyses of major uniparental FGH haplogroups showed no founder effect. The uniparental FGH haplotypes were scattered broadly within the corresponding networks of the ancestral Asian populations, without any FGH cluster identifiable (Additional files 3 and 5). This absence of founder effect has echoed in the results estimating a demographical expansion on the mitochondrial data ( $F_s = -16.276 (\pm 0.000)$ ). It indicates that the FGH gene pool has not been affected by geographical isolation during the different phases of Hmong migration, despite their recent status of “refugees” in French Guiana or their former one of “highlanders” in South-East Asia [47]. It reveals that their isolation is only relative, since they interacted frequently with other communities, notably through their economic role. It is also likely that the Hmong constituted sustainable communities whose numbers were sufficiently large such that their gene pool was not affected by founder effects and/or genetic drift.

The second remarkable feature of the FGH genetic inheritance is that it reveals the maintenance of a predominantly Asian identity right throughout the history of the Diaspora. The three genetic systems analysed in this study revealed a conserved Asian ancestry (Asian ancestry (A):  $A_{mtDNA} = 100\%$ ;  $A_{NRY} = 99.1\%$ ;  $A_{Gm} = 96.6\%$ ; Tables 1,2,3). Admixture events

with individuals from other continents were obviously rare due to the relatively recent timeframe, but still might have occurred during the European colonisation of Indochina. In fact there was a low, but non-negligible, estimated component of European ancestry ( $A_{\text{mtDNA}}=0\%$ ;  $A_{\text{NRY}}=0.9\%$ ;  $A_{\text{Gm}}=3.4\%$ ; Tables 1,2,3) which might be explained by punctual admixture with French or English during the end of 20th century, but is more probably due to a shared ancestry with Indian groups (unfortunately no shared haplotype can dissociate these two hypotheses (data not shown)) [36, 48].

Despite a history punctuated by many conflicts and wars, which led to the exodus of one Hmong community toward French Guiana, this population has conserved a representative genetic diversity, linking it directly to the Asian continent. The rates of the FGH genetic diversity ( $H_{\text{mtDNA}}=0.91$ ;  $H_{\text{NRY}}=0.96$ ;  $H_{\text{Gm}}=0.44$ ) are within the range of values found in most Asian groups [4, 16, 49]. Among them, some examples are particularly interesting, such as the Hmong-Mien in Thailand ( $H_{\text{mtDNA}}=0.93$ ;  $H_{\text{NRY}}=1.00$ ;  $H_{\text{Gm}}=0.23$ ) [16, 34] or the Miao groups in southern provinces of China ( $H_{\text{mtDNA}}=0.94$ ;  $H_{\text{NRY}}=1.00$ ;  $H_{\text{Gm}}=0.23$ ) [4, 39, 50]. In addition to the homology, these values also reflect the conservation of a major cultural trait, i.e. kinship based on patrilocality and patrilineage [34]. The desire to protect their cultural integrity gave to the Hmong Diaspora a cultural unity [2]. This has ultimately led to the preservation of the Asian genetic background during the different phases of the Hmong Diaspora.

### **A preserved Hmong genetic identity**

Within the complex, overall Asian genetic diversity, the Hmong have their own specific patterns [3]. In preserving their original Asian gene pool, the FGH have clearly maintained these particularities. Each studied genetic system was found to be composed by the wide spectra of haplotypes frequent in South East Asia. The FGH maternal lineage showed high frequencies of haplogroups M\*(33.8%) and B5a (22.5%), that are comparable to those present in the Miao of the Chinese province of Yunnan ( $Fq_{\text{M}^*}=23,1\%$ ;  $Fq_{\text{B5a}}=20.5\%$ ) [4]. Their

paternal lineages are characterized by high frequencies of O3\* (48.5%) and D\* (24.2%) as in the Miao of the Chinese province of Yunnan ( $Fq_{O3^*}=57.1\%$ ;  $Fq_{D^*}=28.6\%$ ) [51]. Finally, most frequent FGH Gm haplotypes are Gm5\*;1,3,.. (71.6%) and Gm21;1,17,.. (22.1%) that are predominant in the Miao of Thailand ( $Fq_{Gm5^*;1,3,..}=73.9\%$ ;  $Fq_{Gm21;1,17,..}=20.7\%$ ) [16].

These empirical observations are statistically attested by pairwise  $F_{st}$  comparisons computed from haplotypic data. Among Asian populations showing significantly low levels of differentiation ( $F_{st}<0.05$ ;  $p\text{-value}<0.01$ ) with FGH, populations located in north Thailand ( $F_{st_{mtDNA}}=0.03$ ,  $F_{st_{NRY}}=0.05$ ,  $F_{st_{Gm}}=0.03$ ; figures 2,3,4) and in the south-eastern Chinese provinces of Yunnan-Guangxi ( $F_{st_{mtDNA}}=0.04\text{--}0.05$ ,  $F_{st_{NRY}}=0.05$ ,  $F_{st_{Gm}}=0.03\text{--}0.04$ ; figures 2,3,4) are the only ones that show this pattern for all studied systems. The FGH reveal the same patrimony present in their homeland, unifying them genetically to the Asian Hmong.

All other significantly low  $F_{st}$  values reflect a surprising sex-biased genetic patrimony composed by a South maternal imprint ( $South\text{-}A_{mtDNA}=58.2\%$ ; 15 South Asian populations on 24 Asian populations with  $F_{st}<0.05$ ; figure 2) and an East paternal imprint ( $East\text{-}A_{NRY}=72.9\%$ ; 2 East Asian populations on 3 Asian populations with  $F_{st}<0.05$ ; figure 3). Two explanations can be evoked. The first is the relative low genotyping resolution of the mtDNA coding region of FGH M\* lineages and the high frequency of homoplastic HVS-I/II sequences in this particular haplogroup. For example FGH lineages classified by their HVSI/II sequences in the M5 haplogroup, which is rare in East Asia, have the same HVS-I sequence as a Chinese Miao individual which has been characterised as a M7 lineage, frequent in China [4]. It may lead to an over-estimation of the South Asian maternal ancestry. However it is noticeable that these FGH M\* lineages are mostly integrated in the South Asian diversity (Additional file 6). Therefore the second explanation could be a dual origin. These two components would be due to the high number of FGH haplotypes belonging to maternal haplogroup M\*, mostly integrated within the South Asian diversity, and paternal haplogroup O3a\*, scattered within

the East Asian diversity (Additional file 6). It could be interpreted as an old maternal patrimony, inherited from Palaeolithic migrations, admixed with a Neolithic paternal gene pool, probably during the Han expansion [4, 49, 52]. It could also be the results of multiple gene flows occurred during the different exodus of the Hmong Diaspora. The admixture of these two gene flows gave to the FGH population their genetic characteristics which are close to those observed in South East Asia (figure 4). Thus, the FGH have conserved the entire genetic diversity of their complex ancestry, giving to this particular example of the Hmong Diaspora its genetic unity.

In South America, gene and/or cultural flows into non-Native populations are frequently documented. This is for instance illustrated by the genetic structure of the Afro-Brazilian *quilombos* [53], or the Creole languages coupled with the woodworking techniques of the Noir Marron from Suriname and French Guiana [6] taken from the neighbouring European-origin and Native American ethnic groups. As far as the FGH population is concerned, a small amount of allochthonous cultural traits have also percolated through the population. For instance, though traditional Hmong clothing is worn on special occasions, younger Hmong usually wear Western style clothing. Furthermore, many Hmong have been converted to Christianity. Lastly, education is attended in French so that many Hmong are also bilingual [5][5][5][5][5][5][5]. On the contrary, the FGH retain cultural traditions from Laos, especially when dealing with the agricultural practices which is favoured by the similarity of the environment. Mating is also restricted within the Hmong community, and for this to happen the FGH maintain tied links with the Hmong established in the United States and France to look for spouses [5].

Therefore, the FGH population managed to integrate the whole Hmong culture within the French Guianese social network, and the entire Hmong genetic inheritance within the wider French Guianese genetic diversity. Our study provides a first insight into this population and



its genetic diversity, but further studies in other Hmong communities across different continents, using comparable genetic markers, would bring a deeper understanding of the genetic dispersal that occurred during the Hmong Diaspora.

## **ACKNOWLEDGEMENTS**

We are grateful to all people who contributed to the collection of blood samples, and especially Etienne Bois. We wish to acknowledge support from the genomic platform of Genopole Toulouse Midi Pyrénées, where mtDNA and NRY sequencing/genotyping work was performed. We would like to thank Laure Tonasso and François-Xavier Ricaut from the AMIS laboratory for helping in Y-SNP typing, as well as Clyde Francks from the Language and Genetics laboratory for discussions and critical reading of the manuscript. This work was financed by the following research grants: Programme Interdisciplinaire CNRS “Amazonie – Analyse, modélisation et ingénierie des systèmes amazoniens” and the “Centre de Coopération Scientifique et Culturelle Europe-Asie EURASIAT”.

## REFERENCES

- [1] J. Schliesinger, Ethnic groups of Thailand, Bangkok White Lotus Co, Ltd (2000).
- [2] N. Tapp, Hmong Diaspora, Encyclopedia of Diasporas: Immigrant and Refugee Cultures Around the World. New York : Springer (2005) 103-113.
- [3] S. Xu, D. Kangwanpong, M. Seielstad, M. Srikummool, J. Kampuansai, L. Jin, Genetic evidence supports linguistic affinity of Mlabri--a hunter-gatherer group in Thailand, BMC Genet 11 (2010) 18.
- [4] B. Wen, H. Li, S. Gao, X. Mao, Y. Gao, F. Li, F. Zhang, Y. He, Y. Dong, Y. Zhang, W. Huang, J. Jin, C. Xiao, D. Lu, R. Chakraborty, B. Su, R. Deka, L. Jin, Genetic structure of Hmong-Mien speaking populations in East Asia as revealed by mtDNA lineages, Mol Biol Evol 22 (2005) 725-734.
- [5] P.F. Clarkin, Hmong resettlement in French Guiana, Hmong Studies Journal 6 (2005) 1-27.
- [6] N. Brucato, P. Tortevoeye, S. Plancoulaine, E. Guitard, A. Sanchez-Mazas, G. Larrouy, A. Gessain, J.M. Dugoujon, The genetic diversity of three peculiar populations descending from the slave trade: Gm study of Noir Marron from French Guiana, C R Biol 332 (2009) 917-926.
- [7] S. Mazieres, E. Guitard, E. Crubezy, J.M. Dugoujon, M.C. Bortolini, S.L. Bonatto, M.H. Hutz, E. Bois, F. Tiouka, G. Larrouy, F.M. Salzano, Uniparental (mtDNA, Y-chromosome) polymorphisms in French Guiana and two related populations--implications for the region's colonization, Ann Hum Genet 72 (2008) 145-156.
- [8] S. Mazieres, A. Sevin, S.M. Callegari-Jacques, E. Crubezy, G. Larrouy, J.M. Dugoujon, F.M. Salzano, Population genetic dynamics in the French Guiana region, Am J Hum Biol 21 (2009) 113-117.
- [9] S. Mazières, S.M. Callegari-Jacques, S.G. Crosetti, J.M. Dugoujon, G. Larrouy, E. Bois, E. Crubezy, Hutz, M. H. , F.M. Salzano, French Guiana Amerindian demographic history as revealed by autosomal and Y-Chromosome STRs, Annals of Human Biology (2010).
- [10] N. Brucato, O. Cassar, L. Tonasso, P. Tortevoeye, F. Migot-Nabias, S. Plancoulaine, E. Guitard, G. Larrouy, A. Gessain, J.M. Dugoujon, The imprint of the Slave Trade in an African American population: mitochondrial DNA, Y chromosome and HTLV-1 analysis in the Noir Marron of French Guiana, BMC Evol Biol 10 (2010) 314.
- [11] T.M. Karafet, F.L. Mendez, M.B. Meilerman, P.A. Underhill, S.L. Zegura, M.F. Hammer, New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree, Genome Res 18 (2008) 830-838.
- [12] M. van Oven, Revision of the mtDNA tree and corresponding haplogroup nomenclature, Proc Natl Acad Sci U S A 107 (2010) E38-39; author reply e40-31.
- [13] M.P. Lefranc, V. Giudicelli, C. Ginestoux, J. Jabado-Michaloud, G. Folch, F. Bellahcene, Y. Wu, E. Gemrot, X. Brochet, J. Lane, L. Regnier, F. Ehrenmann, G. Lefranc, P. Duroux, IMGT, the international ImMunoGeneTics information system, Nucleic Acids Res 37 (2009) D1006-1012.
- [14] M.P. Lefranc, G. Lefranc, Molecular genetics of immunoglobulin allotype expression, in: F. Shakib (ed.), The human IgG subclasses: molecular analysis of structure, function and regulation, Pergammon Press, Oxford, 1990, pp. 43-78.
- [15] J. Oudin, L'allotypie de certains antigenes proteiques du serum, C R Hebd Seances Acad Sci 242 (1956) 2606-2608.
- [16] J.M. Dugoujon, S. Hazout, F. Loirat, B. Mourrieras, B. Crouau-Roy, A. Sanchez-Mazas, GM haplotype diversity of 82 populations over the world suggests a centrifugal model of human migrations, Am J Phys Anthropol 125 (2004) 175-192.

- [17] C. Coudray, A. Olivieri, A. Achilli, M. Pala, M. Melhaoui, M. Cherkaoui, F. El-Chennawi, M. Kossmann, A. Torroni, J.M. Dugoujon, The complex and diversified mitochondrial gene pool of Berber populations, *Ann Hum Genet* 73 (2009) 196-214.
- [18] R.M. Andrews, I. Kubacka, P.F. Chinnery, R.N. Lightowlers, D.M. Turnbull, N. Howell, Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA, *Nat Genet* 23 (1999) 147.
- [19] M. van Oven, M. Kayser, Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation, *Hum Mutat* 30 (2009) E386-394.
- [20] D.M. Behar, M. van Oven, S. Rosset, M. Metspalu, E.L. Loogvali, N.M. Silva, T. Kivisild, A. Torroni, R. Villems, A "Copernican" reassessment of the human mitochondrial DNA tree from its root, *Am J Hum Genet* 90 (2012) 675-684.
- [21] M.F. Hammer, A.B. Spurdle, T. Karafet, M.R. Bonner, E.T. Wood, A. Novelletto, P. Malaspina, R.J. Mitchell, S. Horai, T. Jenkins, S.L. Zegura, The geographic distribution of human Y chromosome variation, *Genetics* 145 (1997) 787-805.
- [22] I.S.o.G. Genealogy (2012). Y-DNA Haplogroup Tree 2012, Version: 7.32, Version, <http://www.isogg.org/tree/>, [29 June 2012]
- [23] L.L. Field, J.M. Dugoujon, Immunoglobulin allotyping (Gm, Km) of GAW5 families, *Genet Epidemiol* 6 (1989) 31-33.
- [24] J.M. Dugoujon, G.G. de Lange, A. Blancher, S. Alie-Daram, Y. Marty, Characterization of an IgG2, G2m(23) anti-Rh-D antibody, *Vox Sang* 57 (1989) 133-136.
- [25] M.P. Giraldo, M. Vallet, E. Guitard, M.T. Senegas, A. Sevin, R.M. Nogues, M.P. Aluja, J.M. Dugoujon, GM and KM immunoglobulin allotypes in a Spanish Pyrenean population: Val d'Aran, *Ann Hum Biol* 25 (1998) 453-465.
- [26] L. Excoffier, G. Laval, S. Schneider, Arlequin (version 3.0): An integrated software package for population genetics data analysis, *Evol Bioinform Online* 1 (2005) 47-50.
- [27] P. Forster, A. Torroni, C. Renfrew, A. Rohl, Phylogenetic star contraction applied to Asian and Papuan mtDNA evolution, *Mol Biol Evol* 18 (2001) 1864-1881.
- [28] T. Polzin, D.S.V., On Steiner trees and minimum spanning trees in hypergraphs, *Operations Research Letters* 32 (2003) 12-20.
- [29] R. Qamar, Q. Ayub, A. Mohyuddin, A. Helgason, K. Mazhar, A. Mansoor, T. Zerjal, C. Tyler-Smith, S.Q. Mehdi, Y-chromosomal DNA variation in Pakistan, *Am J Hum Genet* 70 (2002) 1107-1124.
- [30] I. Dupanloup, G. Bertorelle, Inferring admixture proportions from molecular data: extension to any number of parental populations, *Mol Biol Evol* 18 (2001) 672-675.
- [31] N. Maca-Meyer, M. Arnay, J.C. Rando, C. Flores, A.M. Gonzalez, V.M. Cabrera, J.M. Larruga, Ancient mtDNA analysis and the origin of the Guanches, *Eur J Hum Genet* 12 (2004) 155-162.
- [32] Y.G. Yao, L. Nie, H. Harpending, Y.X. Fu, Z.G. Yuan, Y.P. Zhang, Genetic relationship of Chinese ethnic populations revealed by mtDNA sequence diversity, *Am J Phys Anthropol* 118 (2002) 63-76.
- [33] H. Oota, T. Kitano, F. Jin, I. Yuasa, L. Wang, S. Ueda, N. Saitou, M. Stoneking, Extreme mtDNA homogeneity in continental Asian populations, *Am J Phys Anthropol* 118 (2002) 146-153.
- [34] D. Besaggio, S. Fuselli, M. Srikumool, J. Kampuansai, L. Castri, C. Tyler-Smith, M. Seielstad, D. Kangwanpong, G. Bertorelle, Genetic variation in Northern Thailand Hill Tribes: origins and relationships with social structure and linguistic differences, *BMC Evol Biol* 7 Suppl 2 (2007) S12.
- [35] M. Stoneking, F. Delfin, The human genetic history of East Asia: weaving a complex tapestry, *Curr Biol* 20 (2010) R188-193.

- [36] K. Thangaraj, G. Chaubey, T. Kivisild, D. Selvi Rani, V.K. Singh, T. Ismail, D. Carvalho-Silva, M. Metspalu, L.V. Bhaskar, A.G. Reddy, S. Chandra, V. Pande, B. Prathap Naidu, N. Adarsh, A. Verma, I.A. Jyothi, C.B. Mallick, N. Shrivastava, R. Devasena, B. Kumari, A.K. Singh, S.K. Dwivedi, S. Singh, G. Rao, P. Gupta, V. Sonvane, K. Kumari, A. Basha, K.R. Bhargavi, A. Lalremruata, A.K. Gupta, G. Kaur, K.K. Reddy, A.P. Rao, R. Villems, C. Tyler-Smith, L. Singh, Maternal footprints of Southeast Asians in North India, *Hum Hered* 66 (2008) 1-9.
- [37] Q.P. Kong, C. Sun, H.W. Wang, M. Zhao, W.Z. Wang, L. Zhong, X.D. Hao, H. Pan, S.Y. Wang, Y.T. Cheng, C.L. Zhu, S.F. Wu, L.N. Liu, J.Q. Jin, Y.G. Yao, Y.P. Zhang, Large-scale mtDNA screening reveals a surprising matrilineal complexity in East Asia and its implications to the peopling of the region, *Mol Biol Evol* (2010).
- [38] H. Huang, M. Su, X. Li, H. Li, D. Tian, Y. Gao, Y. Guo, Y-chromosome evidence for common ancestry of three Chinese populations with a high risk of esophageal cancer, *PLoS One* 5 (2010) e11118.
- [39] T. Zhao, D.L. Tsung, Gm and Km allotypes in 74 Chinese populations: a hypothesis of the origin of the Chinese nation, *Human Genetics* 83 (1989) 101-110.
- [40] C. Nunez, M. Baeta, C. Sosa, Y. Casalod, J. Ge, B. Budowle, B. Martinez-Jarreta, Reconstructing the population history of Nicaragua by means of mtDNA, Y-chromosome STRs, and autosomal STR markers, *Am J Phys Anthropol* 143 (2010) 591-600.
- [41] P. Anagnostou, V. Coia, G. Spedini, G. Destro-Bisol, Mitochondrial, Y-chromosomal and autosomal variation in Mbenzele Pygmies from the Central African Republic, *Coll Antropol* 34 (2010) 535-543.
- [42] R.A. Kittles, A.W. Bergen, M. Urbanek, M. Virkkunen, M. Linnoila, D. Goldman, J.C. Long, Autosomal, mitochondrial, and Y chromosome DNA variation in Finland: evidence for a male-specific bottleneck, *Am J Phys Anthropol* 108 (1999) 381-399.
- [43] V. Coia, F. Brisighelli, F. Donati, V. Pascali, I. Boschi, D. Luiselli, C. Battaggia, C. Batini, L. Taglioli, F. Cruciani, G. Paoli, C. Capelli, G. Spedini, G. Destro-Bisol, A multi-perspective view of genetic variation in Cameroon, *Am J Phys Anthropol* 140 (2009) 454-464.
- [44] T.S. Simonson, J. Xing, R. Barrett, E. Jerah, P. Loa, Y. Zhang, W.S. Watkins, D.J. Witherspoon, C.D. Huff, S. Woodward, B. Mowry, L.B. Jorde, Ancestry of the Iban is predominantly Southeast Asian: genetic evidence from autosomal, mitochondrial, and Y chromosomes, *PLoS One* 6 (2011) e16338.
- [45] R. Rubicz, P.E. Melton, V. Spitsyn, G. Sun, R. Deka, M.H. Crawford, Genetic structure of native circumpolar populations based on autosomal, mitochondrial, and Y chromosome DNA markers, *Am J Phys Anthropol* 143 (2010) 62-74.
- [46] M. Coelho, C. Coia, V. Alves, D. Luiselli, A. Useli, T. Hagemeijer, A. Amorim, G. Destro-Bisol, J. Rocha, Human Microevolution and the Atlantic Slave Trade: A Case Study from Sao Tome, *Current Anthropology* 49 (2008) 134-143.
- [47] M. Bodner, B. Zimmermann, A. Rock, A. Kloss-Brandstatter, D. Horst, B. Horst, S. Sengchanh, T. Sanguanersri, J. Horst, T. Kramer, P.M. Schneider, W. Parson, Southeast Asian diversity: first insights into the complex mtDNA structure of Laos, *BMC Evol Biol* 11 (2011) 49.
- [48] M. Metspalu, T. Kivisild, E. Metspalu, J. Parik, G. Hudjashov, K. Kaldma, P. Serk, M. Karmin, D.M. Behar, M.T. Gilbert, P. Endicott, S. Mastana, S.S. Papiha, K. Skorecki, A. Torroni, R. Villems, Most of the extant mtDNA boundaries in south and southwest Asia were likely shaped during the initial settlement of Eurasia by anatomically modern humans, *BMC Genet* 5 (2004) 26.

- [49] K.D. Kwak, H.J. Jin, D.J. Shin, J.M. Kim, L. Roewer, M. Krawczak, C. Tyler-Smith, W. Kim, Y-chromosomal STR haplotypes and their applications to forensic and population studies in east Asia, *Int J Legal Med* 119 (2005) 195-201.
- [50] R.J. Gan, S.L. Pan, L.F. Mustavich, Z.D. Qin, X.Y. Cai, J. Qian, C.W. Liu, J.H. Peng, S.L. Li, J.S. Xu, L. Jin, H. Li, Pinghua population as an exception of Han Chinese's coherent genetic structure, *J Hum Genet* 53 (2008) 303-313.
- [51] H. Zhong, H. Shi, X.B. Qi, Z.Y. Duan, P.P. Tan, L. Jin, B. Su, R.Z. Ma, Extended Y chromosome investigation suggests postglacial migrations of modern humans into East Asia via the northern route, *Mol Biol Evol* 28 (2011) 717-727.
- [52] F. Xue, Y. Wang, S. Xu, F. Zhang, B. Wen, X. Wu, M. Lu, R. Deka, J. Qian, L. Jin, A spatial analysis of genetic structure of human populations in China reveals distinct difference between maternal and paternal lineages, *Eur J Hum Genet* 16 (2008) 705-717.
- [53] A.K. Ribeiro-dos-Santos, J.M. Pereira, M.R. Lobato, B.M. Carvalho, J.F. Guerreiro, S.E. Batista Dos Santos, Dissimilarities in the process of formation of Curiaú, a semi-isolated Afro-Brazilian population of the Amazon region, *Am J Hum Biol* 14 (2002) 440-447.

## Figures and table captions

Figure 1. Geographic location of the two Hmong communities sampled in French Guiana.

Figure 2. Map and level of genetic variance ( $F_{st}$ ) between the FGH and 139 Asian populations (black crosses) examined for the first hypervariable segment of the mtDNA.

Figure 3. Map and level of genetic variance ( $F_{st}$ ) between the FGH and 50 Asian populations (black crosses) examined for the non-recombinant region of the Y-chromosome.

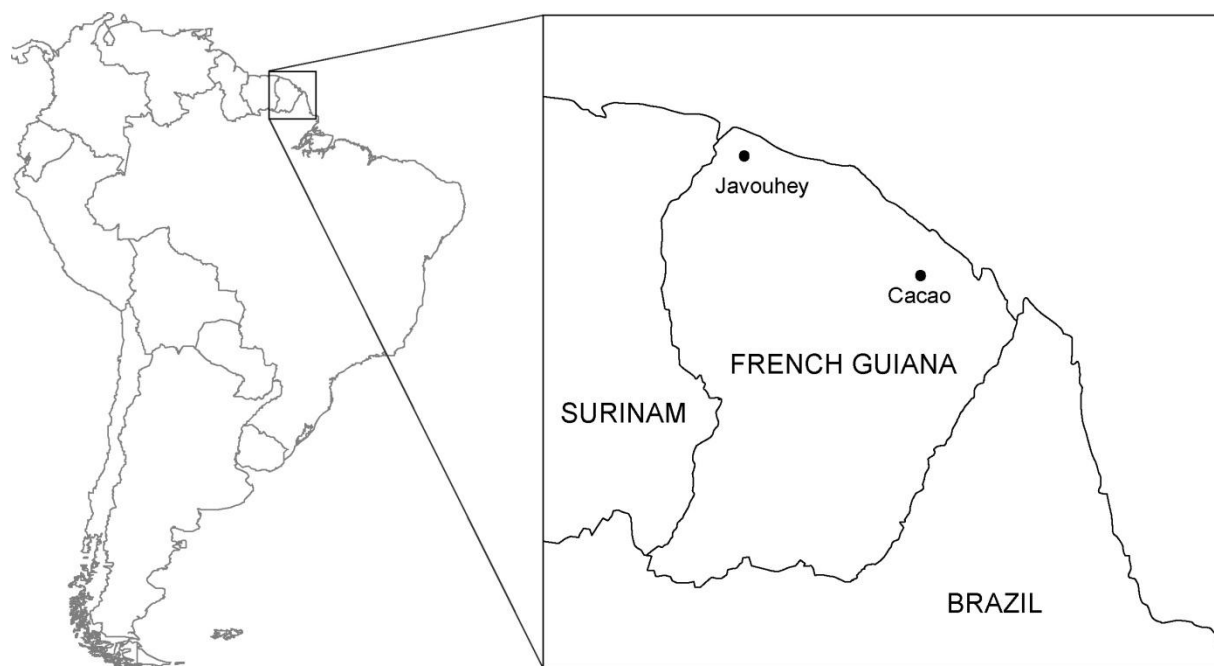
Figure 4. Map and level of genetic variance ( $F_{st}$ ) between the FGH and 173 Asian populations (black crosses) examined for the Gm immunoglobulin system.

Table 1 Estimates of shared maternal ancestries of the FGH. n: number of individuals; LS: Lineages Sharing [31]; *mY*: Dupanloup and Bertorelle's method [30].

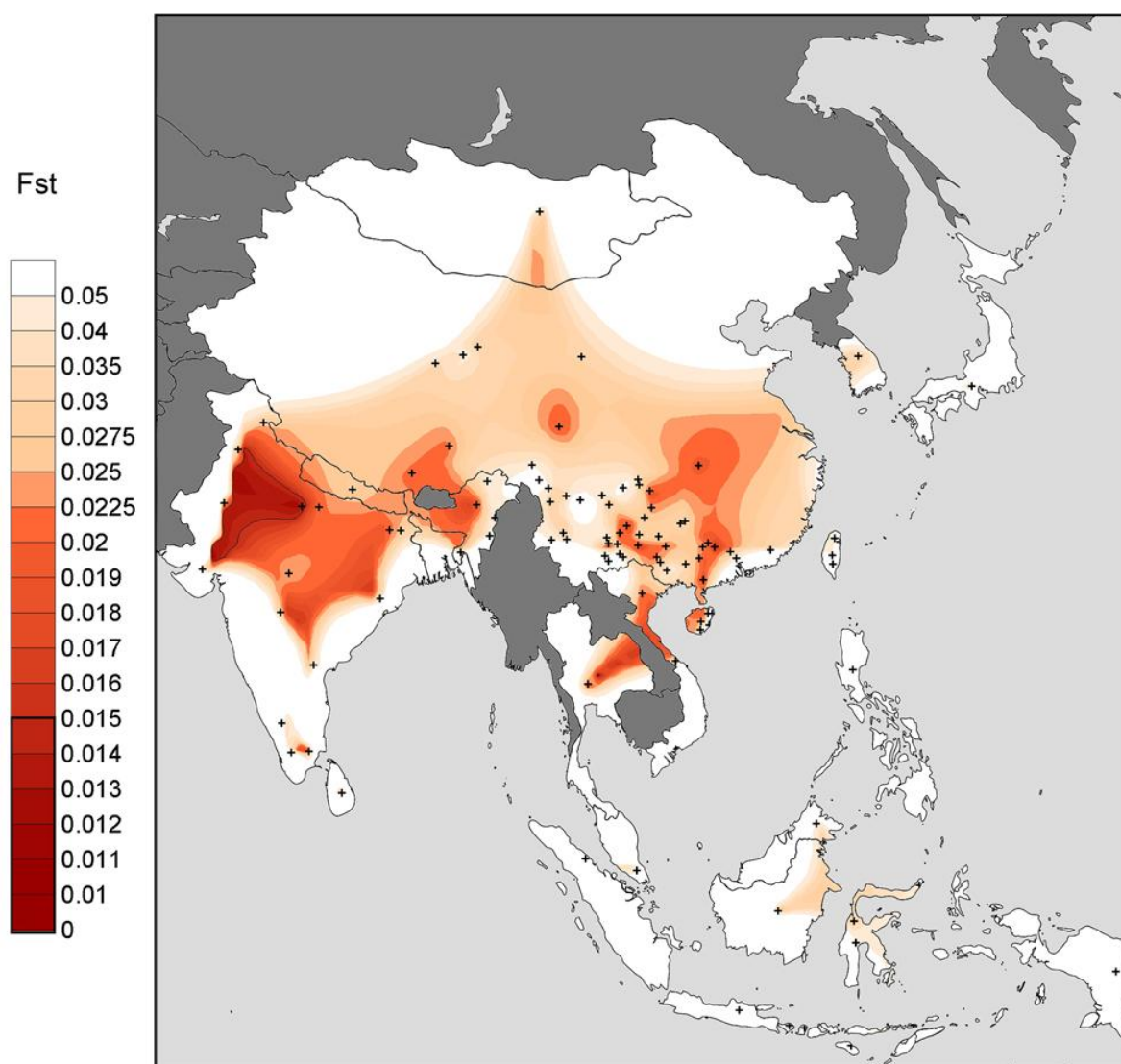
Table 2 Estimates of shared paternal ancestries of the FGH. n: number of individuals; LS: Lineages Sharing [31]; *mY*: Dupanloup and Bertorelle's method [30].

Table 3 Estimators of shared biparental ancestries of the FGH.. n: number of individuals; LS: Lineages Sharing [31]; *mY*: Dupanloup and Bertorelle's method [30].

**Figure 1**

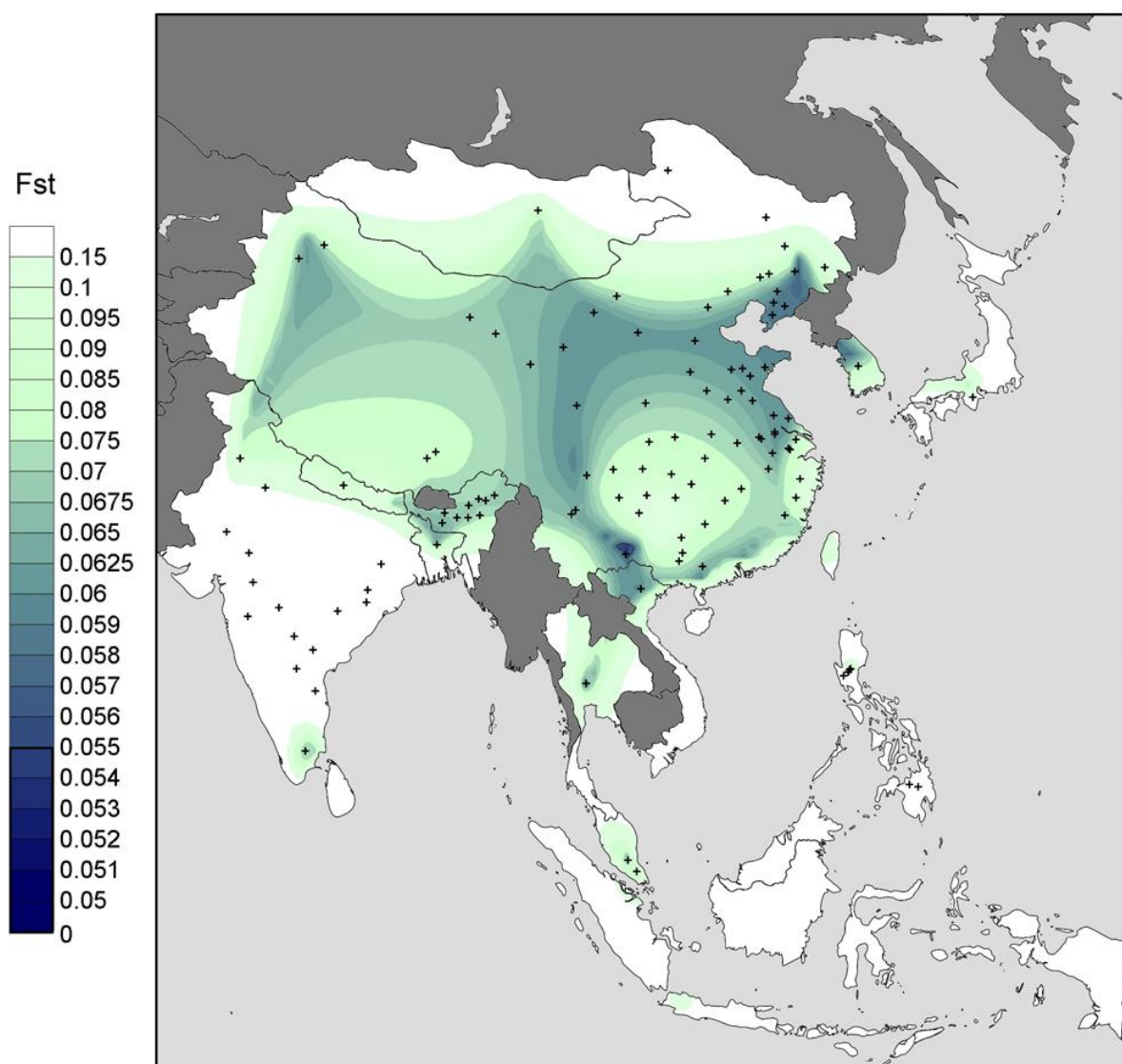


**Figure 2**

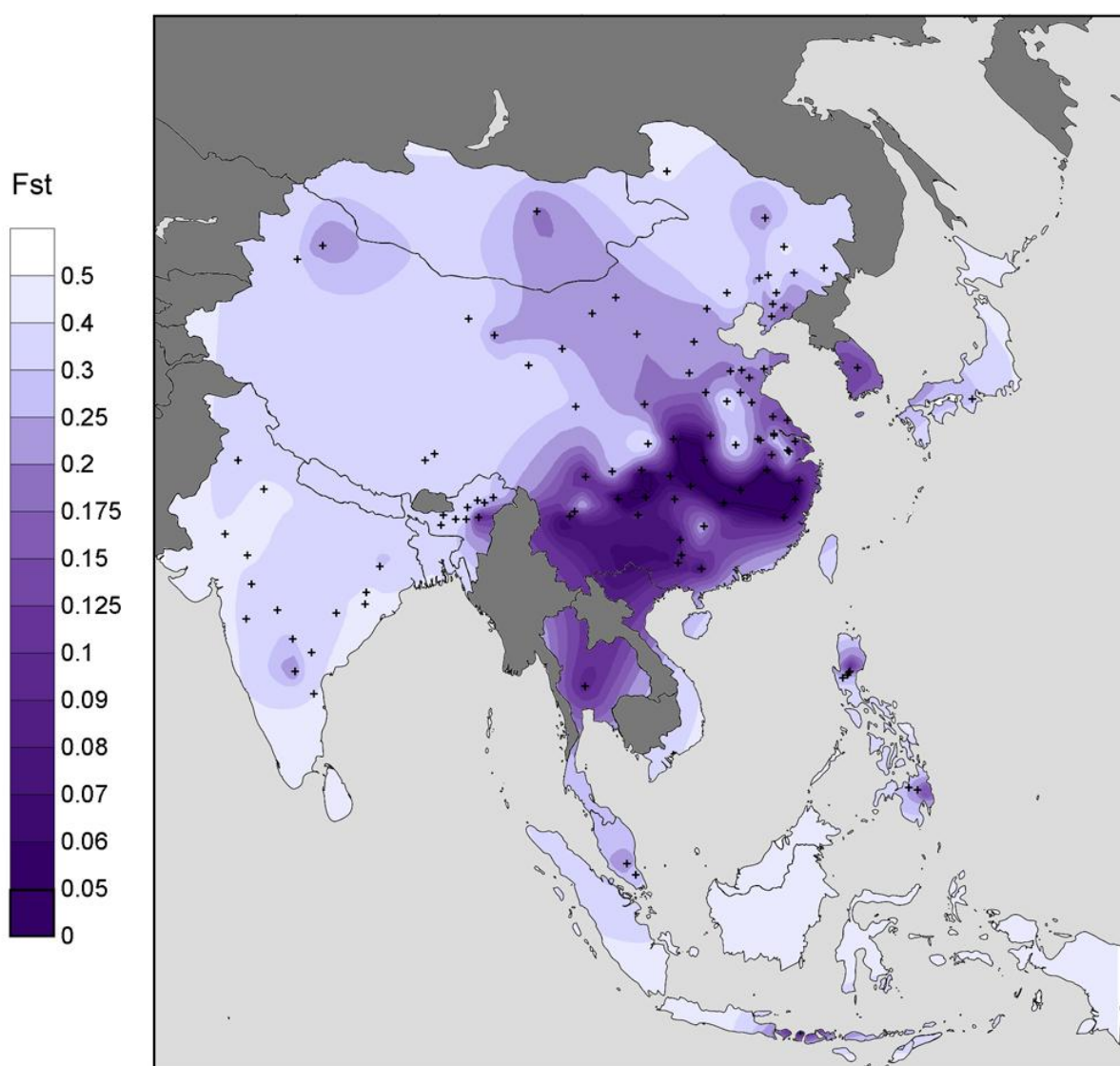




**Figure 3**



**Figure 4**



**Table 1** Estimates of shared maternal ancestries of the FGH. n: number of individuals; LS: Lineages Sharing [31]; *mY*: Dupanloup and Bertorelle's method [30].

Methods	Mainland South East Asia (n=732)	South Asia (n=4157)	East Asia (n=4743)	Insular South East Asia (n=768)	Europe (n=1586)
LS	0.201	0.453	0.203	0.14	0
<i>mY</i>	0.289 ( $\pm 0.085$ )	0.711 ( $\pm 0.085$ )	0 ( $\pm 41.051$ )	0 ( $\pm 1.909$ )	0 ( $\pm 1.805$ )
Average	0.246	0.582	0.102	0.07	0

**Table 2** Estimates of shared paternal ancestries of the FGH. n: number of individuals; LS: Lineages Sharing [31]; *mY*: Dupanloup and Bertorelle's method [30].

Methods	Mainland South East Asia (n=598)	South Asia (n=1196)	East Asia (n=7645)	Insular South East Asia (n=108)	Europe (n=2730)
LS	0.290	0.068	0.459	0.089	0.019
<i>mY</i>	0 ( $\pm 301.036$ )	0 ( $\pm 1133.41$ )	1 ( $\pm 0.519$ )	0 ( $\pm 25.559$ )	0 ( $\pm 0.519$ )
Average	0.145	0.034	0.729	0.045	0.009

**Table 3** Estimators of shared biparental ancestries of the FGH. n: number of individuals; LS: Lineages Sharing [31]; *mY*: Dupanloup and Bertorelle's method [30].

Methods	Mainland South East Asia (n=795)	South Asia (n=2408)	East Asia (n=12917)	Insular South East Asia (n=795)	Europe (n=1204)
LS	0.324	0.161	0.199	0.288	0.026
<i>mY</i>	0.594 ( $\pm 0.402$ )	0 ( $\pm 0.122$ )	0 ( $\pm 0.107$ )	0.363 ( $\pm 0.421$ )	0.042 ( $\pm 0.029$ )
Average	0.459	0.081	0.099	0.326	0.034

## **Supplementary Information**

Additional file 1. References of the populations compiled in the databases.

Additional file 2. MtDNA haplotypes in the French Guianese Hmong sampling.

Additional File 3. Median-joining phylogenetic tree of French Guianese Hmong and Asian mitochondrial HVS-I haplotypes belonging to haplogroup M\*.

Additional file 4. NRY haplotypes in the French Guianese Hmong sampling.

Additional File 5. Reduced-Median and Median-joining phylogenetic tree of French Guianese Hmong and Asian Y-STR haplotypes belonging to haplogroup O3a\*.

Additional file 6. Matching probabilities between female and male FGH haplotypes and the database.